

Using Artificial Neural Network Techniques for Prediction of Electric Energy Consumption

Hasan M. H. Owda · Babatunji Omoniwa
· Ahmad R. Shahid · Sheikh Ziauddin

Received: date / Accepted: date

Abstract Due to imprecision and uncertainties in predicting real world problems, artificial neural network (ANN) techniques have become increasingly useful for modeling and optimization. This paper presents an artificial neural network approach for forecasting electric energy consumption. For effective planning and operation of power systems, optimal forecasting tools are needed for energy operators to maximize profit and also to provide maximum satisfaction to energy consumers. Monthly data for electric energy consumed in the Gaza strip was collected from year 1994 to 2013. Data was trained and the proposed model was validated using 2-Fold and K-Fold cross validation techniques. The model has been tested with actual energy consumption data and yields satisfactory performance.

Keywords Artificial neural network · Electric energy consumption · Modeling · 2-Fold cross validation · K-Fold cross validation

1 Introduction

Numerous techniques for forecasting electric energy consumption have been proposed in the last few decades. For operators, energy consumption (load) forecast is useful in effectively managing power systems. Consumers can also benefit from the forecasted information in order to yield maximum satisfaction. In addition to these economic reasons, load forecasting has also been used for system security purposes. When deployed to handle system security problems, it provides expedient information for detecting vulnerabilities in advance.

H. M. H. Owda · B. Omoniwa · A. R. Shahid · S. Ziauddin
COMSATS Institute of Information Technology, Park Road, Islamabad, Pakistan
Tel.: +92-33-65879709
E-mail: hasanouda@gmail.com

B. Omoniwa
National Mathematical Centre, Abuja, Nigeria
E-mail: tunjiomoniwa@gmail.com

A. R. Shahid
E-mail: ahmadrshahid@comsats.edu.pk

S. Ziauddin
E-mail: sheikh.ziauddin@comsats.edu.pk

Forecasting energy consumed within a particular geographical area greatly depends on several factors, such as, historical load, mean atmospheric temperature, mean relative humidity, Population, GDP per Capita. Over the years, there has been rapid growth annually of about 10% from year 1999 to 2005 for energy demand in the Gaza strip. With about 75% of energy demands from service and household sectors, these demands are barely met [1,2]. In order to meet these demands and efficiently utilize the limited energy, it is imperative to observe historic trends and make futuristic plans based on past data.

In the past, computationally easier approaches like regression and interpolation, have been used, however, this methods may not give sufficiently accurate results. As advances in technology and sophisticated tools are made, complex algorithmic approaches are introduced and more accuracy at the expense of heavy computational burden can be observed. Several algorithms have been proposed by several researchers to tackle electric energy consumption forecasting problem. Previous works can be grouped into three[3]:

Time Series Approach: In this approach, the trend for electric energy consumption is handled as a time series signal. Future consumption is usually predicted based on various time series analysis techniques. However, time series approach is characterized with prediction inaccuracies of prediction and numerical instability. This inaccurate results is due to the fact the approach does not utilize weather information. Studies have shown that there is a strong correlation between the behavior of energy consumed and weather variables. Zhou R. *et al.* [4] proposed a data driven modeling method using time series analysis to predict energy consumed within a building. The model in [4] was applied on two commercial building and is limited to energy prediction within a building. Basu K. *et al.* [5] also used the time series approach to predict appliance usage in a building for just an hour.

Simmhan Y. *et al.* [6] used an incremental time series clustering approach to predict energy consumption. This method in [6] was able to minimize the prediction error, however, very large number of data points were required. Autoregressive integrated moving average (ARIMA) is a vastly used time series approach. ARIMA model was used by Chen J. *et al.* [7] to predict energy consumption in Jiangsu province in China based on data collected from year 1985 to 2007. The model [7] was able to accurately predict the energy consumption, however it was limited to that environment. The Previous works on time series usually use computationally complex matrix-oriented adaptive algorithms which, in most scenarios, may become unstable.

Functional Based Approach: Here, a functional relationship between a load dependent variable (usually weather) and the system load is modelled. Future load is then predicted by inserting the predicted weather information into the pre-defined functional relationship. Most regression methods use functional relationships between weather variables and up-to-date load demands. Linear representations are used as forecasting functions in conventional regression methods and this method finds an appropriate functional relationship between selected weather variables and load demand. Liu D. *et al.* [8] proposed a support vector regression with radial basis function to predict energy consumption in a building. The approach in [8] was only able to forecast the energy consumed due to lighting for some few hours.

In [10], a grey model, multiple regression model and a hybrid of both were used to forecast energy consumption in Zhejiang province of China. Yi W. *et al.* [9] proposed an LS-SVM regression model to also forecast energy consumption. However, these models were limited to a specific geographic area.

Soft Computing Based Approach: This is a more intelligent approach that is extensively being used for demand side management. It includes techniques such as fuzzy logic, genetic algorithm and artificial neural networks (ANN) [11, 12, 13, 14, 15]. The ANN approach is based on examining the relationship that exist between input and output variables. ANN approach was used in [14] to forecast regional load in Taiwan. Empirical data was used to effectively develop an ANN model which was able to predict the regional peak load. Catalão J. P. S. *et al.* [15] used the ANN approach to forecast short-term electricity prices. Levenberg-Marquardt's algorithm was used to train data and the resulting model [15] was able to accurately forecast electricity prices. However, it was only able to predict electricity prices for about 168 hours.

Pinto T. *et al.* [16] also worked on developing an ANN model to forecast electricity market prices with a special feature of dynamism. This model [16] performs well when a small set of data is trained, however, it is likely to perform poorly with large number of data due to the computational complexities involved. Load data from year 2006 to 2009 were gathered and used to develop an ANN model for short-term load forecast in [17]. In [18], ANN Hybrid with Invasive Weed Optimization (IWO) was employed to forecast the electricity prices in the Australian Market. The hybrid model [18] showed good performance, however, the focus was on predicting electricity prices in Australia. Most of the ANN models developed in existing work considered some specific geographic area [14, 18], some models were able to forecast energy consumption for buildings [8, 10, 11] and only for few hours [5, 15].

This study used available historical data¹ from year 1994 to 2013 (but trained data from year 1994 to 2011) in determining a suitable model. The resulting model from training will be used to predict electric energy consumption for future years², while the error criteria such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are used as measures to justify the appropriate model [15, 17, 19]. The model was used to predict the behavior for year 2012 and 2013. The remainder of this paper is organized as follows: Section 2 gives a brief description of the ANN concept. Section 3 presents the ANN approach used to analyse our data. Section 4 evaluates the performance of the ANN model. Section 5 draws conclusions.

2 ANN Concept

ANN is a system based on the working principles of biological neural networks, and is defined as a mimicry of biological neural systems. ANN's are at the vanguard of computational systems designed to create or mimic intelligent behavior. Unlike classical Artificial Intelligence (AI) systems that are aimed at directly emulating rational and logical reasoning, ANN's are targeted at reproducing the causal processing mechanisms that give results to intelligence as a developing property of complex systems.

ANN systems have been designed for fields such as capacity planning, business intelligence, pattern recognition, robotics. . In computer science and engineering, ANN techniques has gained a lot of grounds and is vastly deployed in areas such as forecasting, data analytics and even data mining. The science of raw data examination with the aim of deriving useful conclusions can simply be defined as data analytics. On the other hand, data mining describes the process of determining new patterns from large data sets, by applying a vast set

¹ Data was collected from the Palestinian Bureau of Statistic [1] and the World bank website [2]

² Predictions for 2012 and 2013 were compared with actual data of corresponding years in Section 4

of approaches from statistics, artificial intelligence, or database management. Forecasting is useful in predicting future trends with reliance on past data. The focus on this paper will be on using the ANN approach in forecasting energy consumption. Practically, ANN provides accurate approximation of both linear and non-linear functions.

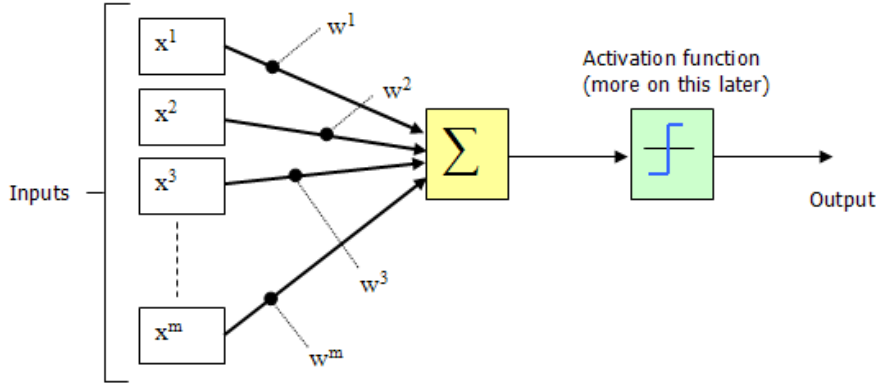


Fig. 1 Mathematical Abstraction of the ANN

A mathematical abstraction of the internal structure of a neuron is shown in Fig. 1. Despite the presence of noise or incomplete information, it is possible for neurons to learn the behaviour or trends and consequently, make useful conclusions. Usually, a neural network is trained to perform a specific function by adjusting weight values between elements as seen in Fig. 2(a). The neural network function is mainly determined by the connections between the elements [17]. By observing the data, it is possible for the ANN to make accurate predictions. Unlike other forecasting approach, the ANN technique has the ability to predict future trends with theoretically poor, but rich data set.

3 Model Process

In this section, we present the overall modeling process. The implementation of the ANN model can be described using the flow chart in Fig. 2(b)³. Historical monthly data (historical energy consumption, mean atmospheric temperature, mean relative humidity, Population, GDP per Capita) has been gathered from years 1994 to 2013 from Gaza region as shown in Fig. 3. Based on the gathered data⁴, this study will develop a forecast model that predicts energy consumption for year 2012 and 2013. Using the ANN technique, training and learning procedures are fundamental in forecasting future events. The training of feed-forward networks is usually carried out in a supervised manner [15]. With a set of data to be trained (usually extracted from the historical data), it is possible to derive an efficient forecast model. The proper selection of inputs for ANN training plays a vital role to the

³ The error criteria used in this study are mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE)

⁴ Data was collected from the Palestinian Bureau of Statistic [1] and the World bank website [2]

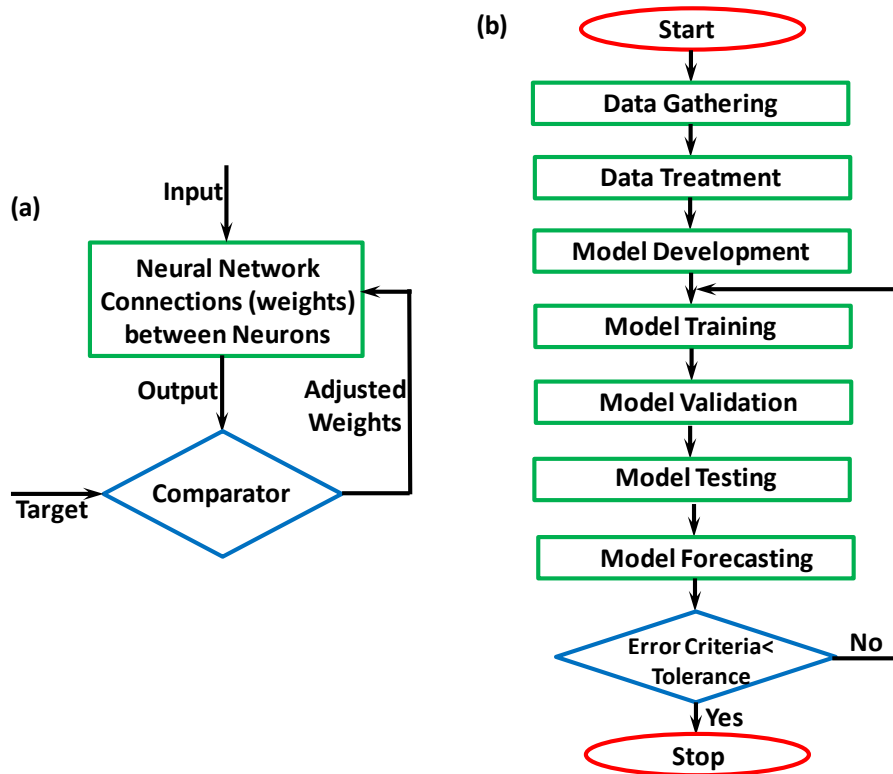


Fig. 2 (a) Neural Network Function (b) Flow Chart of the Model Process

success of the training process. On the other hand, the learning process involves providing both input and output data, the network processes the input and compares the resulting output with desired result. The system then adjusts the weights which acts as a control for error minimization. In order to minimize error, the process is repeated until a satisfactory criterion for convergence is attained. The knowledge acquired by the ANN via the learning process is tested by applying it to a new data set that has not been used before, called the testing set. It should now be possible that the network is able to make generalizations and provide accurate result for new data. Due to insufficient information, some networks do not converge. It is also noteworthy that over-training the ANN can seriously deteriorate forecasts. Also, if the ANN is fed with redundant or inaccurate information, it may destabilize the system.

Training and learning process should be thorough in order to achieve good results. To accurately forecast, it is imperative to consider all possible factors that influence electricity energy consumption, which is not feasible in reality. Electric energy consumption is influenced by a number of factors, which includes: historical energy consumption, mean atmospheric temperature, mean relative humidity, Population, GDP per Capita, PPP, etc. In this paper, different criteria were used to evaluate the accuracy of the ANN approach in forecasting electric energy consumption in Gaza. They include: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).

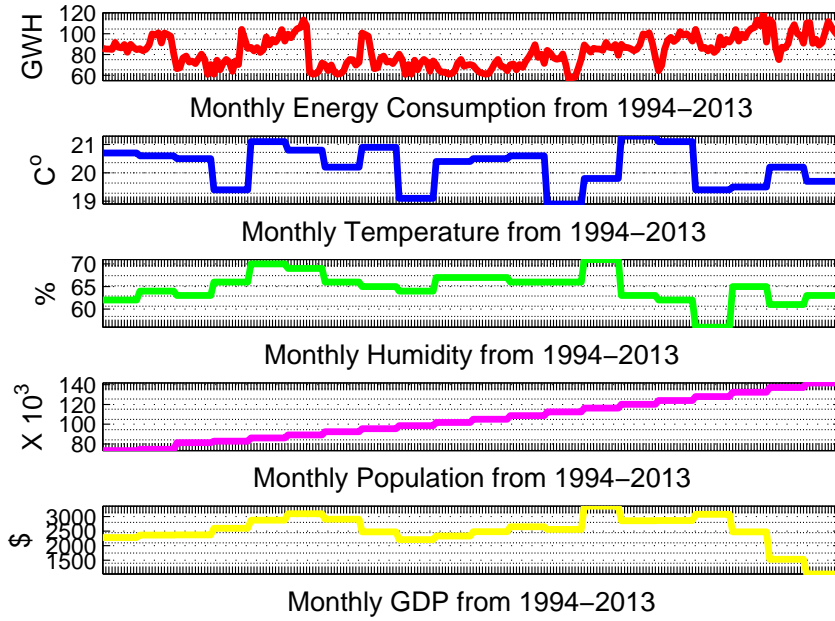


Fig. 3 Historical Data for Gaza from 1994 - 2013

A popular and important criterion used for performance analysis is the MSE. It is used to relay concepts of bias, precision and accuracy in statistical estimation. Here, the difference between the estimated and the actual value is used to get the error, the average of the square of the error gives an expression for MSE. The MSE criterion is expressed in Equation (1).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Where y_i is the actual data and \hat{y}_i is the forecasted data. The RMSE is a quadratic scoring rule which measures the average magnitude of the error. The RMSE criterion is expressed in Equation (2). RMSE usually provides a relatively high weight to large errors due to the fact that averaging is carried out after errors are squared. This makes this criterion an important tool when large errors are specifically undesired.

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

The MAE measures the average error function for forecasts and actual data with polarity elimination. Equation (3) gives the expression for the MAE criterion used. The MAE is a linear score which implies that all the individual differences are weighted equally in the average.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

MAPE, on the other hand, measures the size of the error in percentage (%) terms. It is calculated as the average of the unsigned percentage error. Equation (4) gives the expression for the MAPE criterion used.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| 100\% \quad (4)$$

Validation techniques are employed to tackle fundamental problems in pattern recognition (model selection and performance estimation). In this study, 2-Fold and K-Fold cross validation techniques will be employed and the validation set will only be used as part of training and not part of the test set. The test set will be used to evaluate how well the learning algorithm works as a whole.

4 Performance Evaluation

The forecast model was simulated to obtain results of the energy consumed for year 2012 and 2013 in Gaza. Table 1 compares the actual and the forecasted energy consumption for year 2012. 2-Fold and K-Fold cross validation techniques were used and the performance of the forecast model based on different error criteria is shown in Table 2. Similarly, Table 3 and 4 shows the results for year 2013. The results obtained have good accuracy and shows that the proposed ANN model can be used to predict future trends of electric energy consumption in Gaza.

Months	Actual - 2012	Forecasted - 2012 (2 Fold)	Forecasted - 2012 (k Fold)
January	113.5800	111.7538	119.9518
February	110.5200	109.7213	109.1621
March	85.3500	86.6753	87.8927
April	74.9800	75.2086	80.8937
May	87.1100	88.7946	90.4256
June	86.1000	87.5484	88.8709
July	90.4600	90.0263	90.3327
August	103.1600	102.9062	90.3061
September	108.6600	107.4464	108.2806
October	99.8500	100.7179	90.4634
November	98.5000	99.9205	99.8268
December	90.8400	93.4780	90.8787

Table 1 Forecasted Result for Year 2012

Year 2012	MSE	RMSE	MAE	MAPE
2-Fold Cross Validation	1.21%	1.10%	1.21%	120.91%
K-Fold Cross Validation	3.87%	1.97%	3.87%	386.54%

Table 2 2-Fold and K-Fold Cross Validation for Year 2012

5 Conclusion

In this paper, an ANN model to forecast electric energy consumption in the Gaza strip was presented. To the best of our knowledge, this is the first of it's kind in existing literature.

Months	Actual - 2013	Forecasted - 2013 (2 Fold)	Forecasted - 2013 (k Fold)
January	110.34	109.2616	110.0026
February	110.67	109.3688	110.1573
March	90.776	92.9769	91.2542
April	94.66	97.1215	100.4115
May	88.93	90.8877	89.2735
June	91.443	93.723	94.6202
July	101.82	102.7202	100.2111
August	111.741	109.7194	110.7063
September	106.772	105.0375	108.7167
October	101.92	102.7778	100.2537
November	98.43	100.4408	100.0337
December	92.53	90.4837	96.5896

Table 3 Forecasted Result for Year 2013

Year 2013	MSE	RMSE	MAE	MAPE
2-Fold Cross Validation	1.74%	1.32%	1.74%	173.76%
K-Fold Cross Validation	1.88%	1.37%	1.88%	187.65%

Table 4 2-Fold and K-Fold Cross Validation for Year 2013

Based on the performance evaluation, the error criteria were within tolerable bounds. Empirical results presented in this paper indicates the relevance of the proposed ANN approach in forecasting electric energy consumption. Future works will consider other forecasting techniques.

References

1. Palestinian Central Bureau of Statistics. "Annual statistics", Ramallah – Palestine, Feb., 2014
2. World Bank Data on West Bank and Gaza. "Data", Sourced from: <http://data.worldbank.org/country/west-bank-gaza>, Apr., 2014
3. Park D. C., El-Sharkawi M. A., Marks II R. J., Atlas L. E. and Damborg M. J., (1991) Electric Load Forecasting using an Artificial Neural Network, *IEEE Transactions on Power Engineering*, vol.6, pp. 442-449
4. Zhou R., Pan Y., Huand Z. and Wang Q., (2013) Building Energy Use Prediction Using Time Series Analysis, *IEEE 6th International Conference on Service-Oriented Computing and Applications*, pp. 309-313.
5. Basu K., Debusschere V. and Bacha S., (2012) Appliance Usage Prediction Using a Time Series Based Classification Approach, *IEEE 38th Annual Conference on IEEE Industrial Electronics Society*, pp. 1217-1222.
6. Simmhan Y. and Noor M. U., (2013) Scalable Prediction of Energy Consumption using Incremental Time Series Clustering, *IEEE International Conference on Big Data*, pp. 29-36.
7. Chen J., Huang X., Li J. and Wang G., (2010) Prediction of Energy Consumption in Jiangsu Based on Combination Model, *IEEE 3rd International Conference on Information and Computing*, pp. 109-112.
8. Liu D. and Chen Q., (2013) Prediction of building lighting energy consumption based on support vector regression, *IEEE 9th Asian Control Conference*, pp. 1-5.
9. Yi W. and Ying L., (2009) Applying LS-SVM to Predict Primary Energy Consumption, *IEEE International Conference on E-Product, E-Service and E-Entertainment*, pp. 1-4.
10. Wang Q., Wang X. and Xia F., (2009) Prediction of building lighting energy consumption based on support vector regression, *IEEE 9th Asian Control Conference*, pp. 1-5.
11. Zhang Y. and Chen Q., (2014) Prediction of Building Energy Consumption Based on PSO - RBF Neural Network, *IEEE International Conference on System Science and Engineering*, pp. 60-63.
12. Adika C. O. and Wang L., (2012) Short Term Energy Consumption Prediction Using Bio-Inspired Fuzzy Systems, *IEEE North American Power Symposium*, pp. 1-6.

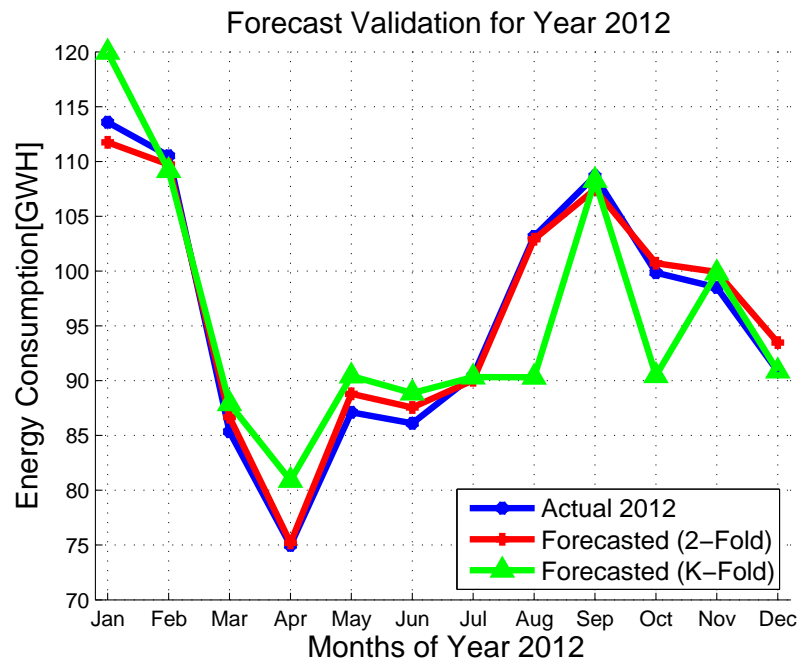


Fig. 4 Plot of Forecast Validation for Year 2012

13. Suganthi L. and Samuel A. A. (2012) Energy models for demand forecasting—A review, *Elsevier Renewable and Sustainable Energy Reviews*, vol. 16(2), pp. 1223–1240
14. Hsu C. C. and Chen C. Y. (2003) Regional load forecasting in Taiwan-applications of artificial neural networks. *Elsevier Energy Conversion and Management*, vol. 44(12), pp. 1941–1949
15. Catalão J. P. S., Mariano S. J. P. S., Mendes V. M. F. and Ferreira L. A. F. M., (2007) An Artificial Neural Network Approach for Short-Term Electricity Prices Forecasting, *IEEE Intelligent Systems Applications to Power Systems*, pp. 1-7.
16. Pinto T., Sousa T. M. and Vale Z., (2012) Dynamic Artificial Neural Network for Electricity Market Prices Forecast, *IEEE 16th International Conference on Intelligent Engineering Systems*, pp. 311-316.
17. Khamis M. F. I., Baharudin Z., Hamid N. H., Abdullah M. F. and Solahuddin S., (2011) Electricity Forecasting For Small Scale Power System Using Artificial Neural Network, *IEEE 5th International Power Engineering and Optimization Conference*, pp. 54-59.
18. Safari M., Dahlan Y. N., Razali N. S. and Rahman T. K., (2013) Electricity Prices Forecasting Using ANN Hybrid with Invasive Weed Optimization (IWO), *IEEE 3rd International Conference on System Engineering and Technology*, pp. 275-280.
19. Kandananond K. (2011) Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach, *Energies*, vol. 4, pp. 1246-1257

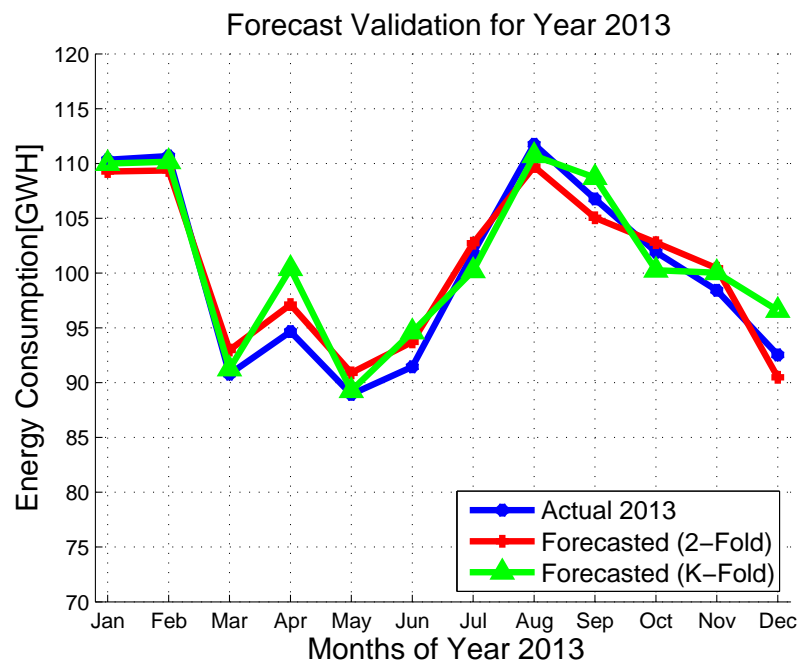


Fig. 5 Plot of Forecast Validation for Year 2013